



ENTERED

OIPE

RAW SEQUENCE LISTING

DATE: 04/24/2002

PATENT APPLICATION: US/09/932,183A

TIME: 15:52:05

Input Set : A:\GC394C1-US-SEQLIST.txt

Output Set: N:\CRF3\04242002\I932183A.raw

```

3 <110> APPLICANT: Estell, David A.
5 <120> TITLE OF INVENTION: Proteases From Gram-Positive Organisms
7 <130> FILE REFERENCE: GC394C1-US
9 <140> CURRENT APPLICATION NUMBER: US 09/932,183A
10 <141> CURRENT FILING DATE: 2001-08-17
12 <150> PRIOR APPLICATION NUMBER: US 09/308,375
13 <151> PRIOR FILING DATE: 1999-05-14
15 <150> PRIOR APPLICATION NUMBER: PCT/US98/18828
16 <151> PRIOR FILING DATE: 1998-09-08
18 <150> PRIOR APPLICATION NUMBER: EP9719636.4
19 <151> PRIOR FILING DATE: 1997-09-15
21 <160> NUMBER OF SEQ ID NOS: 3
23 <170> SOFTWARE: FastSEQ for Windows Version 3.0
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 7100
27 <212> TYPE: DNA
28 <213> ORGANISM: Bacillus subtilis
30 <400> SEQUENCE: 1
31 atattggcat ggtggttatg atgtaattat taagaaagca aacaaagtgc ctcaataact      60
32 gagtggcttt ttcttttgtc ctctccccta ctgaaaggaa gtgattctta cttgagtcaa      120
33 aacctcaaaa ttatactaac cccgcaagct gatacctcat ccaaaactgt cgaacagtta      180
34 aatcagcaaa ttaaatecct ggaaaagaaa ctcaactccc tcaagctcaa tacaacatt      240
35 gattctacaa ccttaaaagc tctgcaagaa ttctcctctg ctatcgacac atatcagaaa      300
36 aacctaaaat cctataatca aacagttaaa gaaacctcaa cagtaattaa gaatgctgac      360
37 ggatcagttg aaaagctcac ccagcagtat aagaaaaatg gtgagatact tcaacgtgaa      420
38 acaaaaataa tcaacaatcg taatacagca ttaaagcaag aaactcaaga ggtaaacaag      480
39 ctaacacagg ccactgagaa actaggacag gttcaaaaaa agactgtgca gagaaatctg      540
40 caaggacagc caacaaaggt agtgcagaaa aaccgccacg ggttcgatga tattgtttat      600
41 acaactgatc ctaaaactaa ttogacctcc tcaaaaacta caactaatta tgaccaacaa      660
42 aggagagcaa ttgagcagct taagcaagat ttagagaagc ttagacagca aggtattggt      720
43 actgatacga ccattctcct tcttgccgca aaaataaaca cagctcaatc cgctcaacaa      780
44 attgaagcac tgcaaaatag gataaggatg ttagatgata aatctgcggc agttgcgaag      840
45 aacaatgaat taaagaaaac cattgaatta tatcagcgac aggcacaagt aaatgttcaa      900
46 aacctaaata cacggtatgg cagttctatg ggctctagta atagacaagc tgttcaagat      960
47 tatttgaaatg cagtaaatag tcttaattgta agcaactggaa gcaataatat cagatcacaa      1020
48 attcaaagct tgaatatgca atttagagaa ttagcctcca acgtcaaac agctgctaatt      1080
49 caagcctctt cttttggagc agaactaacc caaaccttca aaagcatgtc cacctattta      1140
50 atctccggtt cttttattcta cggagctatc tctggaactta aagaaatggt atcccaggca      1200
51 atagaaattg atactctcat gacaaatatt cgccgtggtt tgaatgagcc ggattataaa      1260
52 tataatgaac ttctccaaga atctattgac ttaggtgata cactttcaaa taaaatcaca      1320
53 gatattcttc aaatgacagg cgattttggg agaattgggt tcgatgaaag tgagctctcc      1380
54 acgttaacga aaactgccca agttcttcaa aatgtctctg atttaactcc cgatgatata      1440
55 gttaacactc taacggcagc aatgctcaac tttaatatgg cagcaaatga ttcaatatca      1500

```

RAW SEQUENCE LISTING

DATE: 04/24/2002

PATENT APPLICATION: US/09/932,183A

TIME: 15:52:05

Input Set : A:\GC394C1-US-SEQLIST.txt

Output Set: N:\CRF3\04242002\I932183A.raw

```

56 attgcagata aattaaatga ggttgataat aactatgctg ttacaactct agatctggcc 1560
57 aattctatcc gtaaaagctgg ttcaactgct tctacattcg gggtagagct aaatgatctt 1620
58 attgggtata caactgcaat tgctagtaca acacgtgaat cagggaaat cgtcgggaac 1680
59 tctttaaaga caattttcgc gcggttggg aataatcaaa gctcaattaa agcgttagaa 1740
60 cagattggta tctcagttaa aacagctggg ggtgaagcta aatcagcaag tgatttaatt 1800
61 agtgaagttg ctggttaagt ggatacgctt tctgatgctc agaaacaaaa tacttcaatt 1860
62 ggagtagctg gtatttatca attatcccgt tttaatgcaa tgatgaacaa cttctctatt 1920
63 gctcagaatg cggctaaaaac tgcggctaac tcaacaggaa gtgcttgag tgagcagcaa 1980
64 aagtatgcag atagtctaca agctagggtt aataagcttc aaaataactt cactgaattt 2040
65 gctattgcag cttctgatgc ttttattagc gacggattaa ttgaatttac tcaagccgca 2100
66 ggttctttgc ttaacgcttc tacaggagta atcaaatcag ttgggttcct acctcccctt 2160
67 ttagctgcag taagcactgc aaccctttt ctcagtaaga ataccgcac attagccagc 2220
68 agcctaattt tgggcacacg tgcaatggg caagaaactt tagcgactgc tgggctagaa 2280
69 gctggtatga ctctgcagc agtcgctca agagttctaa aaactgctct tcgagggttg 2340
70 cttgtttcaa ctttagttgg cgggtgcatt gctgctttgg gatggcgct agaataatta 2400
71 atttcttctt ttgcagaagc taaaaaagct aaagatgatt ttgagcagag ccagcaaacc 2460
72 aatgtcgaag caattacgac caataaagac tccactgata aactaataca gcaatataaa 2520
73 gagcttcaaa aagttaaaga gtcaagatct ttaacttcag atgaagagca agaatacctt 2580
74 caagtcactc agcaattagc acaaacttt cctgcattag ttaaaggcta tgattctcaa 2640
75 ggaaatgcaa ttcttaagac aaataaagag cttgaaaaag cgattgagaa tactaaagag 2700
76 tatttggtct taaagaaaca agaaacaaga gacagcgcaa agaaaacatt cgaagacgct 2760
77 tctaaggaaa ttaaaaagtc taaggatgaa ttaaagcagt acaaacaaat agctgactac 2820
78 aacgataaag gtagacctaa atgggatctc attgcagatg acgatgacta taaggttgca 2880
79 gctgataaag ctaaacaaaag tatgctcaaa gctcaatctg acattgagag tggaaatgct 2940
80 aaagttaaag atagcgtcct ttcaattgca aatgcttata gttcaattga ttcagtaata 3000
81 acttttaaaga cgagtattag tgatgttgct aacaaactta acttaaaaga tgatttagat 3060
82 cctgaagaat tagaaaaatt ctctcttctt ttaggaaagc ttcaagaaaa aatgcaaaaa 3120
83 gcttttagatt caggcgatga aaaagctttc gataacgcaa aaaaagatct tcaaagtctc 3180
84 ttggaaacat actccaaatc cgattcttct attgatgttt taaaatgag cttcgacaaa 3240
85 gcacagaaga acataaaaga tggagataag agcttatctt ccgtcaaatc tgaagttggt 3300
86 gatttaggtg agacgctggc agaagcaggt aacgaggcag aagattttgg taagaagcta 3360
87 aaagaagctc tggatgcaaa tagtgttgat gatattaagg cagctattaa agaaatgtca 3420
88 gatgctatgc agttcgatct cgttcaagat gtcttaaatg gggatatttt taataacacc 3480
89 aaagatcaag tagctcctct caatgatctt ctggaaaaaa tggctgaagg taaaagtatt 3540
90 tctgcaaatg aagctaatac ccttattcaa aaagataagg aacttgcca ggctattagc 3600
91 atcgaaaatg gcgttgtaa aattaaccgt gatgaagtta tcaaacaaaag aaaagttaaa 3660
92 cttgatgctt ataacgacat ggttacctac agcaataaat tgatgaaaac agaagttaac 3720
93 aacgctatca aaactttaaa cgtgataacc ttacggattg acagcctgaa aaagctacga 3780
94 aaagaacgaa agcttgatat gtctgaggcc gaactgtcag acctagaagt taagtcaatt 3840
95 aataatgttg cagatgcaaa aaaagaactt aaaaagcttg aagagaaaat gcttcaacct 3900
96 ggtggatact ccaatagtca aattgaagca atgcaaagcg ttaaatcagc tttagaatct 3960
97 tatatttctg catctgaaga agccaccagt acacaagaaa tgaataaaca ggcaattgtt 4020
98 gaagctggaa catcattgga gaattggaca gatcaacaag aaaaagccaa tgaagaaacc 4080
99 aagaacttcca tgtatgttgt tgataaatac aaggaagcat tagaaaaagt taatgctgag 4140
100 attgacaagt acaacaagca ggtcaatgat tatcctaaat actctcagaa atactcgagt 4200
101 gcaatcaaga aagaaattaa agcacttcag caaaagaaaa agcttatgca ggaacaagct 4260
102 aagctgctta aagatcaaat taaatccggt aacattactc aatacggtat tgtaacctct 4320
103 acaacttctt ctggtggaac cccctctca actggtggat catattcagg caagtattca 4380
104 agctacataa attcagcagc tagtaataac aatgttgacc ctgcccttat tgcagctgta 4440

```

RAW SEQUENCE LISTING

DATE: 04/24/2002

PATENT APPLICATION: US/09/932,183A

TIME: 15:52:05

Input Set : A:\GC394C1-US-SEQLIST.txt

Output Set: N:\CRF3\04242002\I932183A.raw

```

105 attcagcaag aatcaggggt taatgctaaa gcacgatctg gtgtaggtgc catgggatta 4500
106 atgcaactga tgccagcaac agcaaaaagc ttaggagtaa ataacgctta cgatccttat 4560
107 caaaatgtta tgggtggaac aaagtacctc gcccaacaac ttgaaaagtt tggcggtaat 4620
108 gttgaaaaag cattggctgc atataatgct gggcctggta acgtaattaa atatggtggt 4680
109 atccctcctt ttaaagaaac acagaattac gtcaagaaga tcatggccaa ctatagcaaa 4740
110 tcgctctcat ctgccacttc ttcaatcgcc agctattata caaataatag cgcttttagg 4800
111 gtaagctoca aatatggaca acaggaatct ggtctccgct cctccccaca caaaggaact 4860
112 gattttgctg caaaagcagg tacagcaatt aaatctcttc aaagtggtaa agtccaaatt 4920
113 gctggctaca gtaaaactgc aggtaaactgg gttgttatta aacaggatga tggaacagtt 4980
114 gccaaagtaca tgcacatgct taacactcct tctgtaaaag caggtcaatc agttaaagcc 5040
115 ggtcaaaacta ttggtaaagt tggtagtaca gggaactcga ctgggaacca ccttcattta 5100
116 cagatcgaac aaaatggaaa aacaatcgat cctgaaaagt acatgcaagg tattggaact 5160
117 tctatttcag atgcgtcaca agctgaggca gaacgacaac aagggatagc tcaggctaaa 5220
118 tctgatcttc tctccctcca aggagatata agttcagtea atgatcagat tcaagaactt 5280
119 cagtatgaac tagttcaatc taaactcgat gagtttgata aaagaatttg agattttgat 5340
120 gttcggatag caaaagatga gtcaatggct aacagataca cttctgacag caaggaattc 5400
121 cgaaaataca cctctgatca gaaaaaagct gtggcagagc aagctaaaat ccaacaacaa 5460
122 aaagttaatt ggattcaaaa agaaattaaa acaataaag cattgaactc cgctcaacgt 5520
123 gcacagcttc aagaagagct taaacaggcc aagctagatt taatttctgt tcaagaccag 5580
124 gttcgtgagc tacagaacaa acttgttcaa tctaaagttg atgagacact taagtcaatt 5640
125 gaaaagtcac cttctaaaaa ccaagggaaa attaaagatg tcgataacaa aatttcaatg 5700
126 actgaagaag atgaagacaa ggttaaatac tatagcaagc aaataaagct cattcaacaa 5760
127 caacaaaagg aagcgaagaa atacattaag cagcttgaag aacaaaagaa agctgcgaaa 5820
128 ggtttccctg acatccagga acagatcact gaagaaatgc aaaactggaa agataaacag 5880
129 aaagatttta accttgagct ttataacacc aagaagtcca tcaaggatat ctataaatca 5940
130 ttggctgatg aagttgtatc catctacaaa gagatgtacg aaaaaatgcg tgatattgag 6000
131 ttagaagcgc atcagaagc gactcaagac ttgatcgatg agatagacaa gactgatgac 6060
132 gaggctaaat ttcaaaaaga attaaaagaa agacaagaca gtattcaaaa gttgactgac 6120
133 caaattaatc aatactctct tgatgattct gaattcggaa agtcaaaagt caaagaacta 6180
134 actgaacagc ttcaaaaaga gcagttagac cttgatgatt ttctaaagga tcgcgaaagt 6240
135 aacaaacgga aagaagcgct ccaagatcag ctcgaaaaag atgaggagtc aatcaacaat 6300
136 aaatacgata atcttgtaaa tgatgaacga gcctttaaaa agcttgagga taagattatg 6360
137 aatggaaaaa tcaccgatat cgctaagcag cttaatgagt tttctaagtt tattaatacc 6420
138 aatatggagt ccattggaaa aagtatttca aacaacctga ttgataaact caaagaagca 6480
139 tctaatagcac tgaatactgc tgtcaaaggc aacacgacag gtaaaaaagt atcctctttc 6540
140 gcttctggag ggtacactgg aacaggatta ggtgctggta aacttgcatt cctacatgac 6600
141 aaagaactga tcttaataaa aactgacaca gccaacatcc ttgatacggg aaaagctggt 6660
142 cgtgaaaccg ctgtggacga ttccccaaaa tggggccaag gagtaaaatt agcagacctt 6720
143 attaaaaaag gaattacttc tattccttca ttagttccta acgttaatca atcaatgtta 6780
144 acaaacagtt taattccaaa tttaaagaag attgagatcc cctcaaaaac aattgcttct 6840
145 tctggagata aaacaattaa tttacgaat actttccaca ttgataagct aataggagga 6900
146 gaatcgggag cgagatcgat gtttgaaagc attaaaaacg aagttgtaaa actaaatggt 6960
147 agcatgtaag agtctgcaaa agcagactct ttatttaact taacttgagg tggaaactca 7020
148 tgattagaga aagtcaatac tttatgttca ataatatccc ttcttatgaa ttaggagccg 7080
149 taaatgtaaa tacagaagga
151 <210> SEQ ID NO: 2
153 <211> LENGTH: 2285
154 <212> TYPE: PRT
155 <213> ORGANISM: Bacillus subtilis

```

RAW SEQUENCE LISTING

DATE: 04/24/2002

PATENT APPLICATION: US/09/932,183A

TIME: 15:52:05

Input Set : A:\GC394C1-US-SEQLIST.txt

Output Set: N:\CRF3\04242002\I932183A.raw

157 <400> SEQUENCE: 2

```

158 Leu Ser Gln Asn Leu Lys Ile Ile Leu Thr Pro Gln Ala Asp Thr Ser
159   1           5           10           15
160 Ser Lys Thr Val Glu Gln Leu Asn Gln Gln Ile Lys Ser Leu Glu Lys
161           20           25           30
162 Lys Leu Asn Ser Leu Lys Leu Asn Thr Asn Ile Asp Ser Thr Thr Leu
163           35           40           45
164 Lys Ala Leu Gln Glu Phe Ser Ser Ala Ile Asp Thr Tyr Gln Lys Asn
165           50           55           60
166 Leu Lys Ser Tyr Asn Gln Thr Val Lys Glu Thr Ser Thr Val Ile Lys
167           65           70           75           80
168 Asn Ala Asp Gly Ser Val Glu Lys Leu Thr Gln Gln Tyr Lys Lys Asn
169           85           90           95
170 Gly Glu Ile Leu Gln Arg Glu Thr Lys Ile Ile Asn Asn Arg Asn Thr
171           100          105          110
172 Ala Leu Lys Gln Glu Thr Gln Glu Val Asn Lys Leu Thr Gln Ala Thr
173           115          120          125
174 Glu Lys Ser Leu Gly Gln Val Gln Lys Lys Thr Val Gln Arg Asn Leu Gln
175           130          135          140
176 Gly Gln Pro Thr Lys Val Val Gln Lys Asn Arg His Gly Phe Asp Asp
177          145          150          155          160
178 Ile Val Tyr Thr Thr Asp Pro Lys Thr Asn Ser Thr Ser Ser Lys Thr
179           165          170          175
180 Thr Thr Asn Tyr Asp Gln Gln Arg Arg Ala Ile Glu Gln Leu Lys Gln
181           180          185          190
182 Asp Leu Glu Lys Leu Arg Gln Gln Gly Ile Val Thr Asp Thr Thr Ile
183           195          200          205
184 Ser Ser Leu Gly Arg Lys Ile Asn Thr Ala Gln Ser Ala Gln Gln Ile
185          210          215          220
186 Glu Ala Leu Gln Asn Arg Ile Arg Met Leu Asp Asp Lys Ser Ala Ala
187          225          230          235          240
188 Val Ala Lys Asn Asn Glu Leu Lys Lys Thr Ile Glu Leu Tyr Gln Arg
189           245          250          255
190 Gln Ala Gln Val Asn Val Gln Asn Leu Asn Thr Arg Tyr Gly Ser Ser
191           260          265          270
192 Met Gly Ser Ser Asn Arg Gln Ala Val Gln Asp Tyr Leu Asn Ala Val
193           275          280          285
194 Asn Ser Leu Asn Val Ser Thr Gly Ser Asn Asn Ile Arg Ser Gln Ile
195          290          295          300
196 Gln Ser Leu Asn Met Gln Phe Arg Glu Leu Ala Ser Asn Ala Gln Thr
197          305          310          315          320
198 Ala Ala Asn Gln Ala Ser Ser Phe Gly Ala Glu Leu Thr Gln Thr Phe
199           325          330          335
200 Lys Ser Met Ser Thr Tyr Leu Ile Ser Gly Ser Leu Phe Tyr Gly Ala
201           340          345          350
202 Ile Ser Gly Leu Lys Glu Met Val Ser Gln Ala Ile Glu Ile Asp Thr
203           355          360          365
204 Leu Met Thr Asn Ile Arg Arg Val Met Asn Glu Pro Asp Tyr Lys Tyr
205          370          375          380

```

RAW SEQUENCE LISTING

DATE: 04/24/2002

PATENT APPLICATION: US/09/932,183A

TIME: 15:52:05

Input Set : A:\GC394C1-US-SEQLIST.txt

Output Set: N:\CRF3\04242002\I932183A.raw

```

206 Asn Glu Leu Leu Gln Glu Ser Ile Asp Leu Gly Asp Thr Leu Ser Asn
207 385                      390                      395                      400
208 Lys Ile Thr Asp Ile Leu Gln Met Thr Gly Asp Phe Gly Arg Met Gly
209                      405                      410                      415
210 Phe Asp Glu Ser Glu Leu Ser Thr Leu Thr Lys Thr Ala Gln Val Leu
211                      420                      425                      430
212 Gln Asn Val Ser Asp Leu Thr Pro Asp Asp Thr Val Asn Thr Leu Thr
213                      435                      440                      445
214 Ala Ala Met Leu Asn Phe Asn Ile Ala Ala Asn Asp Ser Ile Ser Ile
215                      450                      455                      460
216 Ala Asp Lys Leu Asn Glu Val Asp Asn Asn Tyr Ala Val Thr Thr Leu
217 465                      470                      475                      480
218 Asp Leu Ala Asn Ser Ile Arg Lys Ala Gly Ser Thr Ala Ser Thr Phe
219                      485                      490                      495
220 Gly Val Glu Leu Asn Asp Leu Ile Gly Tyr Thr Thr Ala Ile Ala Ser
221                      500                      505                      510
222 Thr Thr Arg Glu Ser Gly Asn Ile Val Gly Asn Ser Leu Lys Thr Ile
223                      515                      520                      525
224 Phe Ala Arg Ile Gly Asn Asn Gln Ser Ser Ile Lys Ala Leu Glu Gln
225                      530                      535                      540
226 Ile Gly Ile Ser Val Lys Thr Ala Gly Gly Glu Ala Lys Ser Ala Ser
227 545                      550                      555                      560
228 Asp Leu Ile Ser Glu Val Ala Gly Lys Trp Asp Thr Leu Ser Asp Ala
229                      565                      570                      575
230 Gln Lys Gln Asn Thr Ser Ile Gly Val Ala Gly Ile Tyr Gln Leu Ser
231                      580                      585                      590
232 Arg Phe Asn Ala Met Met Asn Asn Phe Ser Ile Ala Gln Asn Ala Ala
233                      595                      600                      605
234 Lys Thr Ala Ala Asn Ser Thr Gly Ser Ala Trp Ser Glu Gln Gln Lys
235                      610                      615                      620
236 Tyr Ala Asp Ser Leu Gln Ala Arg Val Asn Lys Leu Gln Asn Asn Phe
237 625                      630                      635                      640
238 Thr Glu Phe Ala Ile Ala Ala Ser Asp Ala Phe Ile Ser Asp Gly Leu
239                      645                      650                      655
240 Ile Glu Phe Thr Gln Ala Ala Gly Ser Leu Leu Asn Ala Ser Thr Gly
241                      660                      665                      670
242 Val Ile Lys Ser Val Gly Phe Leu Pro Pro Leu Leu Ala Ala Val Ser
243                      675                      680                      685
244 Thr Ala Thr Leu Leu Leu Ser Lys Asn Thr Arg Thr Leu Ala Ser Ser
245                      690                      695                      700
246 Leu Ile Leu Gly Thr Arg Ala Met Gly Gln Glu Thr Leu Ala Thr Ala
247 705                      710                      715                      720
248 Gly Leu Glu Ala Gly Met Thr Arg Ala Ala Val Ala Ser Arg Val Leu
249                      725                      730                      735
250 Lys Thr Ala Leu Arg Gly Leu Leu Val Ser Thr Leu Val Gly Gly Ala
251                      740                      745                      750
252 Phe Ala Ala Leu Gly Trp Ala Leu Glu Ser Leu Ile Ser Ser Phe Ala
253                      755                      760                      765
254 Glu Ala Lys Lys Ala Lys Asp Asp Phe Glu Gln Ser Gln Gln Thr Asn

```

VERIFICATION SUMMARY

DATE: 04/24/2002

PATENT APPLICATION: US/09/932,183A

TIME: 15:52:06

Input Set : A:\GC394C1-US-SEQLIST.txt

Output Set: N:\CRF3\04242002\I932183A.raw